

Integración de tecnologías eye-tracking con vídeos interactivos para neurorrehabilitación cognitiva

JM. Martínez-Moreno^{1,2}, J. Solana Sánchez^{1,2}, C. Gómez Pérez³, C. Cáceres Taladriz^{1,2},
S. González Palmero⁴, M. Morell Vilaseca³, T. Roig Rovira³,
J.M. Tormos Muñoz³, E.J. Gómez Aguilera^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación (UPM), Madrid, España
{jmartinez,jsolana,,ccaceres,egomez}@gbt.tfo.upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España.

³ Instituto Universitario de Neurorrehabilitación Guttmann-UAB, Barcelona, España,
{cgomez,mmorell,troig,jmtormos}@guttmann.com

⁴ Lavinia Interactiva, Barcelona, España, samuel.gonzalez@laviniainteractiva.com

Resumen

La integración de las nuevas tecnologías en el proceso de rehabilitación permite la generación de terapias personalizadas, ubicuas y basadas en la evidencia. Tecnologías como el vídeo interactivo son propicias para el desarrollo de entornos virtuales en los que el paciente se ve inmerso dentro de actividades de la vida diaria en los que tiene que lograr un objetivo ecológico en un contexto seguro, controlado y adaptado a su perfil disfuncional. Dentro de este marco de rehabilitación la interacción visual paciente-entorno virtual se entiende como el mecanismo de comunicación principal, siendo además la atención visual un reflejo del estado cognitivo del paciente.

El trabajo presentado en este artículo permite la integración de un sistema de eye-tracking con un entorno de neurorrehabilitación basado en vídeo interactivo. El objetivo último del sistema es la monitorización en tiempo real de la atención visual del usuario durante el proceso de neurorrehabilitación. Esta monitorización permite no sólo reproducir la ejecución de la actividad junto con el foco de la mirada, sino también detectar faltas de atención por parte del usuario, que permiten al vídeo interactivo reaccionar y adaptar la presentación de estímulos para ayudar a centrar su atención y así completar el objetivo de la actividad.

1. Introducción

El deterioro cognitivo es la principal causa de discapacidad de las sociedades desarrolladas [1][2]. En el informe de la OMS sobre repercusión global de las enfermedades, 3 de las 5 primeras causas en los países desarrollados asocian alteraciones cognitivas que por sí solas, o agravando la discapacidad física, generan una disminución importante de la capacidad productiva (sobre las personas y sus familiares) y un elevado consumo de recursos.

La rehabilitación neuropsicológica y, dentro de ésta, la estimulación cognitiva, permite recuperar algunas de las funciones alteradas o prevenir el deterioro asociado al envejecimiento, aprovechando la naturaleza plástica del sistema nervioso, optimizando su capacidad de reorganización funcional y estimulando la creación de nuevos patrones de activación. A pesar de la existencia de un núcleo importante de conocimiento empírico sobre la

bondad de la rehabilitación neuropsicológica, existen limitaciones que dificultan su extensión a la mayor parte de beneficiarios potenciales: el modelo clásico de intervención presencial no es asumible (por falta de profesionales especializados, y por la intensidad, duración y personalización necesaria); y la falta de guías de práctica clínica que permitan una extensión racional de estos servicios.

Durante los últimos años se están aplicando importantes avances tecnológicos en rehabilitación, evolucionando desde sistemas tradicionales basados en sesiones presenciales de trabajo personal y síncrono entre terapeuta y paciente, hasta la implantación de nuevos sistemas de rehabilitación que faciliten la personalización de las intervenciones, modulen la intensidad y duración de los programas, monitoricen en tiempo real y diferido y permitan obtener conocimiento derivado de la práctica clínica y su estructuración [3].

Los Entornos Virtuales (EV) interactivos permiten simular situaciones reales situando al usuario dentro de un escenario enriquecido. Explotar las ventajas de los EVs para el tratamiento neuropsicológico permite la exposición fuera de riesgo del paciente ante situaciones de la vida cotidiana y dentro de un entorno controlado, ya que permite graduar la dificultad de la situación en función de los estímulos, distractores y elementos de ayuda que se ofrezcan [4].

El Vídeo Interactivo (VI) es una tecnología que permite al terapeuta trabajar con escenarios virtuales donde recrear las actividades de la vida diaria más importantes para cada individuo y donde la demanda ambiental para cada sujeto se ajuste al nivel de trabajo y la modulación pautada por el terapeuta. En este trabajo de investigación se entiende el vídeo interactivo como cualquier aplicación de vídeo en la que la secuenciación y selección de los mensajes (información) se determina por la respuesta del usuario.

Como se mostró en el trabajo previo [5] el uso de la tecnología de VI es viable como elemento base en la rehabilitación cognitiva. El empleo de esta tecnología en

dicho campo de investigación, permite la generación de entornos virtuales basados en interacciones sencillas, intuitivas y de fácil asimilación debido en parte a la utilización de imágenes reales.

El funcionamiento del sistema visual humano al interactuar con EVs se encuentra estrechamente ligado al estado cognitivo de los pacientes, el cual es representativo de las lesiones cerebrales que puedan haber sufrido [6][7]. Capturar y analizar la información referente al comportamiento visual de un paciente puede ayudar a conocer la eficacia de las tareas de rehabilitación. Los dispositivos de eye-tracking [8] permiten realizar la monitorización del comportamiento visual, ofreciendo información precisa acerca de las fijaciones visuales realizadas por el paciente. Gracias a esta información es posible conocer el foco de la atención de un paciente y su estado cognitivo.

Las técnicas de eye-tracking tienen un gran potencial de aplicación en una amplia variedad de disciplinas y áreas de estudio, desde el marketing y la publicidad, para conocer cómo y qué capta la atención de los clientes, hasta la investigación médica o la psicolingüística. Además de en la aplicación clínica, se emplea en sistemas de comunicación aumentados y adaptado, pasando también por los estudios de usabilidad. En esta última clase de estudios, el objetivo es determinar la forma en la que el usuario explora visualmente la interfaz a través de la que interactúa con el sistema.

El uso de esta tecnología permite, no sólo monitorizar y registrar el foco de la mirada durante la ejecución de una determinada actividad, sino también utilizar la información capturada como medio de interacción. Así, por ejemplo, podremos saber qué llama su atención o qué le distrae, si ha visto un determinado estímulo o no, si ha prestado atención a las instrucciones recibidas o si por el contrario ha realizado la tarea sin tener conciencia de lo que debe hacer y, por tanto, el resultado registrado en la actividad puede ser engañoso.

Además, reproduciendo la ejecución de una actividad junto con el registro de la mirada del paciente, el terapeuta será capaz de evaluar el resultado de una actividad concreta de manera más objetiva, ya que sabría en cada momento a qué parte de la pantalla dirige su atención el paciente.

Asimismo, el registro de la atención visual realizado en tiempo real, permite determinar si el paciente está mirando donde debería para alcanzar el objetivo final marcado en la actividad. Detectando estas faltas de atención, se pueden definir alertas y mensajes que ayuden a centrar su atención, consiguiendo así completar la tarea pautada.

2. Materiales y metodología

2.1. Vídeo Interactivo

El vídeo interactivo se presenta no solo como una tecnología para mostrar información de forma dinámica sino como un medio que involucra al usuario de forma activa dentro del proceso de comunicación.

Para ello, se asocian interacciones a cualquier elemento de los que aparecen en una escena del vídeo, quedando definidas así las zonas calientes del mismo. De esta forma, durante la ejecución del vídeo el usuario podrá interactuar con dichas regiones, modificando de este modo el flujo de acción de la tarea.

Una de las características más relevantes de los EVs generados mediante VI es que están generados sobre imagen real enriquecida. Esto se realiza incorporando a la imagen real elementos virtuales (textos, imágenes, audios, etc.) mediante los cuales se consigue graduar el nivel de exigencia requerido por el entorno y adaptarse así al esquema cognitivo del paciente.

2.2. El Entorno Virtual

Con el objetivo de validar las hipótesis de partida que sitúan al VI como tecnología propicia para la generación de EVs para la rehabilitación cognitiva, se elaboró una prueba de concepto presentada en el trabajo [5] y cuyo desarrollo se ha evolucionado.

El resultado de este demostrador es un EV interactivo en el que el paciente se enfrentará a la actividad de vida diaria “ir a comprar el pan”. Durante su ejecución, se recorrerá de forma virtual una serie de escenas (desde la inicial en que se encuentra sentado en el sofá de casa, hasta la final en la que compra el pan en la panadería) necesarias para alcanzar el objetivo final. El paciente participa de forma activa en el escenario, ya que será él quien a través de sus interacciones establezca el flujo de acción entre las diferentes escenas.

En la prueba de concepto desarrollada, el primer escenario sitúa al usuario en la casa, donde encontrará un entorno familiar que le recuerde su propio día a día. Después, en el escenario de la calle, Figura 1, el usuario encontrará un cúmulo de distractores y personajes con los que interactuar. Finalmente en la panadería se enfrentará a una situación en la cual debe interactuar con los personajes que se encuentra para conseguir, por el camino adecuado, el objetivo que le ocupa.



Figura 1. Escena del entorno virtual interactivo

La actividad se desarrolla dentro de un entorno completamente controlado siendo el terapeuta quien, mediante la definición de los estímulos mostrados, gradúa el nivel de exigencia de las habilidades cognitivas que se le requerirá al paciente durante su ejecución. También, y de la misma forma, define ayudas que recibirá el paciente para guiarle a través de la tarea y que de una forma libre de errores alcance el objetivo de la misma.

El guión de la actividad ha sido desarrollado por el equipo clínico del Institut Guttmann. Además de la descripción de la actividad también se establecieron en él los indicadores más relevantes que se monitorizarán durante la ejecución del paciente. En la AVD desarrollada los parámetros almacenados serán el número de aciertos, omisiones y errores, y el tiempo de ejecución, así como un seguimiento de la atención visual del paciente mediante la tecnología de eye-tracking.

2.3. Eye-tracker Tobii 1750

El dispositivo de eye-tracking utilizado es el Tobii 1750 [9], que viene integrado en una pantalla de 17", como se puede ver en la Figura 2.



Figura 2. Eye-tracker Tobii modelo 1750

La información que captura el dispositivo es la siguiente:

- Componentes vertical y horizontal del foco de la mirada para cada ojo. Es decir, la posición de la pantalla donde mira el usuario, representada como una coordenada (x,y), donde el punto (0,0) equivale a la esquina superior izquierda, y el (1,1) a la esquina inferior derecha.
- Distancia de cada ojo, que nos indica cuánto de alejado se encuentra el usuario respecto de la pantalla.
- Instante de tiempo, indicado en microsegundos, en que se ha realizado la captura de datos. La diferencia de tiempo entre muestras equivale a la frecuencia de muestreo.
- Diámetro de cada una de las pupilas del usuario.
- Código de validez para las medidas realizadas sobre cada ojo, que nos permite saber la precisión de la medida realizada.

Dentro del Sistema Visual Humano [8], el movimiento de los ojos tiene dos estados: las sacadas, que son movimientos rápidos del ojo (alrededor de 30 milisegundos) que el ser humano utiliza para centrar el objetivo dentro del área de enfoque; y las fijaciones, que ocurren cuando el movimiento se para, permitiendo al ojo la adquisición de contenido.

A la hora de analizar los datos capturados sobre la mirada, antes que el propio conjunto de datos completo de seguimiento de la mirada, suelen ser de mayor interés la posición espacial y duración de las fijaciones, así como

los instantes en que se producen los movimientos sacádicos.

3. Resultados

El principal resultado del trabajo presentado es la integración del dispositivo de eye-tracking Tobii 1750 con el VI, generando un entorno de monitorización e interacción del usuario a través de su atención visual.

Para ello, el mayor reto abarcado en el desarrollo ha sido el de conseguir utilizar las librerías del sistema Tobii desde la aplicación web del VI. Para ello, se ha implementado un servlet que corre sobre Tomcat, que bajo petición del vídeo interactivo, hace las llamadas a las librerías del Tobii, pudiendo después enviar la información capturada por el eye-tracking de manera que pueda ser utilizada por el vídeo interactivo.

Para que las librerías del Tobii puedan ser accedidas desde el servlet se ha utilizado JNI (Java Native Interface). Este componente hace de puente entre el código C de las librerías nativas del Tobii y el código Java del servlet, que escucha las peticiones provenientes del VI. La introducción intermedia del servlet Tomcat permite de manera rápida y eficaz mandar las acciones necesarias al Tobii, y recoger los datos que éste devuelve.

Para ello, se ha implementado un protocolo de comunicación entre la aplicación web del VI y el servlet en el que se han definido las siguientes acciones:

- **setup:** el VI solicita al servlet que inicie una nueva sesión para un usuario (UID) y actividad dadas (AID); el servlet devuelve un identificador de sesión (SID).
- **start:** el VI solicita al servlet que comience la captura de datos para la sesión actual (SID).
- **query:** solicita saber si el usuario está prestando o no atención a la actividad; el servlet devuelve un booleano que lo indica.
- **stop:** solicita al servlet que detenga la captura de datos temporalmente, bien porque se ha terminado la actividad o bien porque se ha detenido momentáneamente (*pause*).
- **end:** solicita al servlet que termine la sesión, y que almacene todos los datos capturados en un fichero; el servlet devuelve la ruta del fichero donde se han guardado los datos para que la aplicación de VI pueda reproducir la actividad junto con la posición de la mirada en cada momento.

Para el envío de la información desde el servlet hacia el VI con la información pertinente en cada acción descrita anteriormente, se ha utilizado JSON (JavaScript Object Notation), ampliamente utilizada en aplicaciones web y que permite de manera sencilla y ligera el intercambio de información estructurada. En la Figura 3 se muestra el esquema de integración utilizado, donde se ve cómo el servlet hace de puente entre el VI y el eye-tracker.

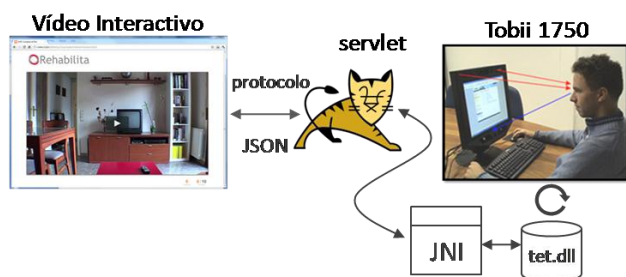


Figura 3. Esquema de integración

Así pues, en primer lugar el entorno de monitorización permite registrar la atención visual del usuario mientras realiza una actividad de VI. Este registro, almacenado en un fichero, permite posteriormente al VI reproducir la secuencia de ejecución de la actividad junto con la información del foco de la mirada, según fue realizada por el usuario. Esto aporta al terapeuta información extra para evaluar la forma en que el paciente ha ejecutado la tarea pautada.

En segundo lugar, el entorno también es capaz de detectar faltas de atención del usuario (como son cerrar los ojos o mirar fuera de la pantalla), lo que permite al VI adaptarse a dichas situaciones, intentando llamar la atención del usuario mediante alertas sonoras y textuales. Asimismo, estas alertas posibilitan al VI modificar el flujo de acción de la tarea, para ayudar al paciente a completar el objetivo definido en la actividad.

4. Conclusiones

El entorno de monitorización visual presentado en este trabajo permite el registro del foco de la atención visual del usuario dentro de un entorno virtual elaborado mediante la tecnología de Vídeo Interactivo (VI), en el que el paciente ejecutará las actividades de vida diaria pautadas como parte del proceso de neurorrehabilitación cognitiva.

Este registro de la atención sirve, por un lado, para aportar a los terapeutas información adicional sobre cómo ha ejecutado la actividad el usuario, de manera que se reproduce dónde ha dirigido su mirada en cada instante de la ejecución de la actividad. Esta información permitirá evaluar de manera más objetiva el resultado concreto obtenido por el paciente, sabiendo qué estímulos ha percibido y cuáles no, y dónde estaba mirando a la hora de realizar cada interacción.

El otro gran logro es ser capaces de detectar faltas de atención, durante la ejecución de la tarea. Estas alertas detectadas también podrán utilizarse como entrada para modificar el flujo de acción de la tarea, adaptándonos en tiempo real al comportamiento del paciente. Además, la detección de estas faltas de atención nos permite lanzar alertas, mediante mensajes textuales y sonoros, cuya intención es llamar la atención del usuario y centrarla en los estímulos adecuados para completar el objetivo terapéutico.

Agradecimientos

Este trabajo de investigación ha sido parcialmente financiado por el proyecto CENIT-E “REHABILITA” CEN-20091043. Los autores expresan su agradecimiento a todos los socios del consorcio REHABILITA.

Referencias

- [1] C.J. Murray, A.D. Lopez. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997a; 349: 1498-504.
- [2] World Health Organization. Burden of Disease Statistics; Available online: <http://www.who.org/> (último acceso agosto de 2012).
- [3] Tormos JM et Al. REHABILITA: Tecnologías Disruptivas para la Rehabilitación del Futuro. *Actas CASEIB 2010: XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*.
- [4] Gamito P, Oliveira J, Morais D, Rosa P, Saraiva T. NeuAR – A Review of the VR/AR Applications in the Neuroscience Domain. Andrew Yeh Ching Nee. *Augmented Reality-Some Emerging Application Areas*. 2011. pp 131 – 154. ISBN: 978-953-307-422-1
- [5] Martínez-Moreno JM, Sánchez-González P et Al. Entornos virtuales de vídeo interactivo para neurorrehabilitación cognitiva. *Actas CASEIB 2011 XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*.
- [6] Suh M, Basu S, Kolster R, Sarkar R, McCandliss B, Ghajar J. Increased oculomotor deficits during target blanking as an indicator of mild traumatic brain injury. *Neuroscience Letters*, 410(3), 203-207. 2006.
- [7] Suh M, Kolster, RA, Niogi S, McCandliss B, Ivry RB, Voss HU, Sarkar R, Ghajar J. Degree of brain connectivity predicts eye-tracking variability. *Journal of Korean Physical Society*, 53, 3468-3473. 2008.
- [8] Duchowski A. *Eye tracking methodology: theory and practice*. 2nd edition. Springer. 2007
- [9] Tobii, <http://www.tobii.com>. Último acceso, agosto 2012.